GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 00:49:36 ; Search time 4784.11 Seconds
                 (without alignments)
                 1569.471 Million cell updates/sec

Title: US-09-698-781-3
Perfect score: 1436
Sequence: 1 MKQILHPALETTRAMTLFPVL.........KHQLVRDSCKASCNCSNSIY 258

Scoring table: BLOSUM62
                 Xgapop   10.0 , Xgapext   0.5
                 Ygapop   10.0 , Ygapext   0.5
                 Fgapop   6.0 , Fgapext   7.0
                 Delop   6.0 , Delext   7.0

Searched: 2054640 seqs, 1455102878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0969878l/runat_07032003_083459_5322/app_query.fasta_1.654
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cd1 -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000000
-USER=US0969878l_@CGN_1_1_2566_@runat_07032003_083459_5322 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   GenEmbl:*
                 1: gb_ba:*   29: em_vi:*
                 2: gb_htg:*   30: em_htg_hum:*
                 3: gb_in:*   31: em_htg_inv:*
                 4: gb_om:*   32: em_htg_other:*
                 5: gb_ov:*   33: em_htg_mus:*
                 6: gb_pat:*   34: em_htg_pln:*
                 7: gb_ph:*   35: em_htg_rod:*
                 8: gb_pl:*   36: em_htg_mam:*
                 9: gb_pr:*   37: em_htg_vrt:*
                 10: gb_ro:*   38: em_sy:*
                 11: gb_sts:*   39: em_htgo_hum:*
                 12: gb_sy:*   40: em_htgo_mus:*
                 13: gb_un:*   41: em_htgo_other:*
                 14: gb_vi:*
                 15: gb_ba:*
                 16: gb_fun:*
                 17: em_hum:*
                 18: em_in:*
                 19: em_mu:*
                 20: em_om:*
                 21: em_or:*
                 22: em_ov:*
                 23: em_pat:*
                 24: em_ph:*
                 25: em_pl:*
                 26: em_ro:*
                 27: em_sts:*
                 28: em_un:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	100.0	2138	9	HSSPG28	X94323 H.sapiens m
2	1436	100.0	2144	6	AX127587	AX127587 Sequence
3	1386	96.5	2128	6	AX335634	AX335634 Sequence
4	1386	96.5	2128	9	HSCRISP3G	X95240 H.sapiens m
5	1335	69.8	1335	9	HUMTPX1A	M25532 Human testi
6	1002	69.8	1380	9	BC022011	BC022011 Homo sapi
7	1002	69.8	1406	9	HSCRISP2I	X95339 H.sapiens m
8	982.5	68.4	1388	10	CPU35712	U35712 Cavia porce
9	952.5	66.3	1295	4	ECCRISP3	AJ001400 Equus cab
10	856	59.6	1280	10	AB009662	AB009662 Rattus no
11	853	59.4	1432	10	AF078552	AF078552 Rattus no
12	835	58.1	1418	10	MUSTPX1A	M25533 Mouse testl
13	835	58.1	1380	10	MUSAEG1A	M92849 Mouse acidi
14	772.5	53.8	1403	10	MUSCRISPA	L05559 Mouse cyste
15	760.5	53.0	1445	10	BC011150	BC011150 Mus muscu
16	759.5	52.9	928	10	RNESGR	X04643 Rat mRNA fo
17	759.5	52.9	1498	10	RATAEG	M31173 Rat epididy
18	684	47.6	495	4	ECA6632	AJ006632 Equus cab
19	677.5	47.2	1090	5	U13619	U13619 Heloderma h
20	669	46.6	1336	5	AF384218	AF384218 Agkistrod
21	656	45.7	791	9	AY093955	AY093955 Rhabdophi
22	650.5	45.3	1321	5	AF384219	AF384219 Trimeresu
23	650	45.1	1318	5	U59447	U59447 Trimeresuru
24	637.5	44.4	1305	5	AF384220	AF384220 Laticauda
25	627.5	43.7	1309	5	AF190861	AF190861 Lapemis h
26	625.5	43.6	1394	4	ECA6631	ECA6631 Equus cab
27	625.5	43.6	1406	10	MUSCRISPB	L05560 Mouse cyste
28	625.5	43.6	1418	10	BC022573	BC022573 Mus muscu
29	612.5	42.7	1316	9	AF159541	AF159541 Lapemis h
30	590.5	41.1	453	4	BTA277708	BTA277708 Bos tauru
31	542.5	37.8	1343	9	AF123894	AF123894 Macaca mu
32	542	37.7	1475	4	ECA315379	ECA315379 Equus cab
33	529.5	36.9	1281	9	HUMARP	HUMARP Homo sapien
34	529.5	36.9	1282	9	S80310	S80310 acidic epid
35	529.5	36.9	1886	9	HSCRISP1G	X95237 H.sapiens m
36	525	36.6	1174	4	AJ006631	AJ006631 Equus cab
37	463	32.2	334	4	BTA277709	AJ277709 Bos tauru
38	406.5	28.3	1797	2	HSCRISP1D	X95238 H.sapiens m
39	366	25.5	151752	2	AC010779	AC010779 Homo sapi
40	358	24.9	758	9	AF393653	AF393653 Xenopus l
41	358	24.9	90901	9	HSJ417L20	AL121974 Human DNA
42	350.5	24.4	1025	10	BC025083	BC025083 Mus muscu
43	345	24.0	237	4	SSU66635	AJ006635 Sus scrof
44	337.5	23.5	1008	9	AX376552	AX376552 Sequence
45	329	22.9	873	9	AF400440	AF400440 Homo sapi

RESULT 1   ALIGNMENTS

HSSPG28

HSSPG28	2138 bp mRNA linear PRI 04-MAR-1996
LOCUS	HSSPG28
DEFINITION	H.sapiens mRNA for SGP28 protein.
ACCESSION	X94323
VERSION	X94323.1 GI:1213612
KEYWORDS	glycoprotein; SGP28 protein.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 2138)
AUTHORS	Kjeldsen,L., Cowland,J.B., Johnsen,A.H. and Borregaard,N.
TITLE	SGP28, a novel matrix glycoprotein in specific granules of human neutrophils with similarity to a family of mammalian secretory glycoproteins and a rodent sperm-coating glycoprotein
JOURNAL	FEBS Lett. 380 (3), 246-250 (1996)
MEDLINE	96166934
PUBMED	8601434
REFERENCE	2 (bases 1 to 2138)
AUTHORS	Cowland,J.B.
TITLE	Direct Submission
JOURNAL	Submitted (19-DEC-1995) J.B. Cowland, Granulocyte Research Laboratory, Dept of Hematology, National Univ. Hosp., Rigshospitalet L-4041, 9 Blegdamsvej, 2100 Copenhagen, DENMARK
FEATURES	Location/Qualifiers
source	1..2138

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="bone marrow"

CDS	41..778

/codon_start=1
/product="SGP28 protein"
/protein_id="CAA63984.1"
/db_xref="GI:1213613"
/db_xref="SWISS-PROT:P54108"
/translation="MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVN
KHNELRRAVSPPARNMLKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYM
SSAPSSWSQAIQSWFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQ
KVLKYYVVCQYCPAGNWANRLYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKS
LKLTLCKHQLVRDSCKASCNCSNSIY"

ORIGIN

US-09-698-781-3 (1-258) x HSSPG28 (1-2138)

Alignment Scores:

Pred. No.:	1		
Score:	3.02e-136	Length:	2138
	1436.00	Matches:	258
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

```
Qy      1 MetLysGlnIleLeuHisProAlaLeuGluThrThrThrAlaMetThrLeuPheProValLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 ATGAAACAAATACTTCATCCTGCTCTGGAAACCACTGCAATGACACATTATTCCCAGTGCTG 61

Qy     21 LeuPheLeuValAlaGlyLeuLeuProSerPhePheProAlaAsnGluAspLysAspProAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 TTGTTCCTGGTTGCTGGGCTGCTTCCATCTTTCCAGCAAATGAAGATAAGGATCCCGGCT 121

Qy     41 PheThrAlaLeuLeuThrThrGlnThrGlnValGlnArgGluIleValAsnLysHisAsn 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TTTACTGCTTTGTTAACCACCCAAACACAAGTGCAAAGGGAGATTGTGAATAAGCACAAT 181

Qy     61 GluLeuArgArgAlaValSerProProAlaArgAsnMetLeuLysMetGluTrpAsnLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 GAACTGAGGAGAGCAGTATCTCCCCCGCCAGAAAACATGCTGAAGATGGAATGGAACAAA 241

Qy     81 GluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnGlnCysAsnTyrArgHisSerAsnPro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 GAGGCTGCAGCAAATGCCCAAAAAGTGGGCAAACCAGTGCAATTACAGACACAGTAACCCA 301
```

RESULT 2
AX127587

LOCUS	AX127587 2144 bp DNA linear PAT 15-MAY-2001
DEFINITION	Sequence 2 from Patent WO0131343.
ACCESSION	AX127587
VERSION	AX127587.1 GI:14134279
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 2144)
AUTHORS	Hubert,R.S., Raitano,A.B., Afar,D.E., Mitchell,S.C., Faris,M. and Jakobovits,A.
TITLE	Diagnosis and therapy of cancer using sgp28-related molecules
JOURNAL	Patent: WO 0131343-A 2 03-MAY-2001; Urogenesys, Inc. (US)
FEATURES	Location/Qualifiers
source	1..2144

/organism="Homo sapiens"
/db_xref="taxon:9606"

ORIGIN

US-09-698-781-3 (1-258) x AX127587 (1-2144)

Alignment Scores:

Pred. No.:	1		
Score:	3.02e-136	Length:	2144
	1436.00	Matches:	258
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

```
Qy    101 LysAspProAlaPheThrAlaLeuLeuThrThrGlnThrGlnValGlnArgGluIleValAsn 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 AAGGATCCAGCATTCACAGCTCTGCTCACTACTCAAACTCAAGTGCAGCGAGAAATCGTCAAC 361

Qy    121 SerTrpSerGlnAlaIleGlnSerTrpPheAspGluTyrAsnAspPheAspPheGlyVal 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 TCATGGTCACAAGCAAATCCAAACAGCTGGTTTGATGAGTACAATGATTTTGACTTTGGTGTA 421

Qy    141 GlyProLysThrProAsnAlaValValGlyHisTyrThrGlnValValTrpTyrSer 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 GGGCCAAAGACTCCCAACGCAGTGGTTGGACATTATACACAGGTTGTTTGGTACTCTTCA 481

Qy    161 TyrLeuValGlyCysGlyAsnAlaTyrCysProAsnGlnLysValLeuLysTyrTyr 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 TACCTCGTTGGATGTGGAAATGCCTACTGTCCAAATCAAAAAGTTCTAAAATACTACTAT 541

Qy    181 ValCysGlnTyrCysProAlaGlyAsnTrpAlaAsnArgLeuTyrValProTyrGluGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 GTTTGCCAATATTGTCCTGCTGGTAATTGGGCTAATAGACTAATATTGTCCCTTATGAACAA 601

Qy    201 GlyAlaProCysAlaSerCysProAspAsnCysAspAspGlyLeuCysThrAsnGlyCys 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 GGAGCACCTTGTGCCAGTTGCCCAGATAACTGTGACGATGGACTATGCACCAATGGTTGC 661

Qy    221 LysTyrGluAspLeuTyrSerAsnCysLysSerLeuLysLeuThrLeuCysLysHis 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 AAGTACGAAGATCTCTATAGTAACTGTAAAAGTTTGAAGCTCACATTAACCTGTAAACAT 721

Qy    241 GlnLeuValArgAspSerCysLysAlaSerCysAsnCysSerAsnSerIleTyr 258
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    722 CAGTTGGTCAGGGACAGTTGCAAGGCCTCCTGCAATTGTTCAAACAGCATTTAT 775
```

US-09-698-781-3 (1-258) x AX127587 (1-2144)

```
Qy      1 MetLysGlnIleLeuHisProAlaLeuGluThrThrThrAlaMetThrLeuPheProValLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 ATGAAACAAATACTTCATCCTGCTCTGGAAACCACTGCAATGACACATTATTCCCAGTGCTG 62

Qy     21 LeuPheLeuValAlaGlyLeuLeuProSerPhePheProAlaAsnGluAspLysAspProAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TTGTTCCTGGTTGCTGGGCTGCTTCCATCTTTCCAGCAAATGAAGATAAGGATCCCGGCT 122
```

```
Qy      41 PheThrAlaLeuLeuThrThrGlnThrGlnValGlnArgGluIleValAsnLysHisAsn  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 TTTACTGCTTTGTTAACCACCCAAACACAAGTGCAAAGGGAGATTGAATAAGCACAAT 182

Qy      61 GluLeuArgArgAlaValSerProProAlaArgAsnMetLeuLysMetGluTrpAsnLys  80
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 GAACTGAGGAGGGCAGTATCTCCCCCTGCCAGAAACATGCTGAAGATGGAATGGAACAAA 242

Qy      81 GluAlaIalaAlaAsnAlaGlnLysTrpAlaAsnGlnCysAsnTyrArgHisSerAsnPro 100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GAGGCTGCAGCAAAATGCCCAAAAGTGGGCAAAACATGGTGAATTACAGACACACAACCA 302

Qy     101 LysAspArgMetThrSerLeuLysCysGlyGluAsnLeuTyrMetSerSerAlaProSer 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AAGGATCGAATGACAAGTCTAAAATGTGGTGAGAATCTCTACATGTCAAGTGCCCCCAGC 362

Qy     121 SerTrpSerGlnAlaIleGlnSerTrpPheAspGluTyrAsnAspPheAspPheGlyVal 140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 TCATGGTCACAAGCAAATCCAAAGCTGGTTTGATGAGTACAATGATTTTGGTGTA 422

Qy     141 GlyProLysThrProAsnAlaValValGlyHisThrTyrThrGlnValValPheTyrSer 160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GGGCCAAAGACTCCCAACGCCGGTTGGACATTATACACAGGTTGTTGTTCTACTCTTCA 482

Qy     161 TyrLeuValGlyCysGlyAsnAlaTyrCysProAsnGlnLysValLysTyrYrTyr 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 TACTTCGTTGGATGTGGAAATGCCTACTGTCCCAATCAAAAGTTCTAAAATACTACTAT 542

Qy     181 ValCysGlnThrTyrCysProAlaGlyAsnArgLeuTyrValProTyrGluGln 200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 GTTTGCCAATATTGTCCCTGCTGGTAATAGACTATATGTCCCTTATGGAACAA 602

Qy     201 GlyAlaProCysAlaSerCysProAspAsnCysAspAspGlyLeuCysThrAsnGlyLys 220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 GGAGCACCTTGTCCCAGTCCCAGATAACTGTGACGADGGACTATGCACCAATGGTTGC 662

Qy     221 LysTyrGluAspLeuTyrSerAsnCysLysSerSerLeuLysSerLeuLysSerTyrLys 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 AAGTACGAAGATCTCTATAGTAATAACCTGTAAAAGTTTGAAGCTCACATTAAACCAT 722

Qy     241 GlnLeuValIleArgAspSerCysLysSerAlaSerCysAsnCysSerSerAsnSerIleTyr 258
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 CAGTTGGTCAGGGACAGTTGCAAGGCATCCTGCAATTGTTCAAACAGCATTTAT 776
```

```
Qy       9 LeuGluThrThrAlaMetThrLeuPheProValLeuLeuPheLeuValAlaGlyLeuLeu  28
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTGGAAACCACTGCAATGACATTATTCCCAGTGCTGCTGTTTCTGGTTGCTGGCTGCTT  60

Qy      29 ProSerPheProAlaAsnGluAspLysAspLysProAlaAlaPheThrAlaLeuLeuThrGln  48
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCATCTTTTCCAGCAAATGAAGATAAGGATCCCGCTTTTACTGCTTTGTTAACCACCCAA 120

Qy      49 ThrGlnValGlnArgGluIleValAsnLysHisAsnGluLeuArgArgAlaValSerPro  68
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ACACAAGTGCAAAGGGAGATTGTGAATAAGCACAATGAACTGAGGAGGGCAGTATCTCC 180

Qy      69 ProAlaArgAsnMetLeuLysMetGluTrpAsnLysGluAlaIalaAlaAsnAlaGlnLys  88
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTGCCAGAAACATGCTGAAGATGGAATGGAACAAAGAGGCTGCAGCAAATGCCCAAAAG 240

Qy      89 TrpAlaAsnGlnCysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLys 108
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGGGCAAAACATGGTGAATTACAGACACACAACCAAAGGATCGAATGACAAGTCTAAA 300

Qy     109 CysGlyGluAsnLeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSer 128
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGTGGTGAGAATCTCTACATGTCAAGTGCCCCCAGCTCATGGTCACAAGCAAATCCAAAGC 360

Qy     129 TrpPheAspGluTyrAsnAspPheAspPheGlyValGlyProLysThrProAsnAlaVal 148
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGGTTTGATGAGTACAATGATTTTGGTGTAGGGCCAAAGACTCCCAACGCCGGTG 420

Qy     149 ValGlyHisThrTyrThrGlnValValPheTyrSerTyrLeuValGlyCysGlyAsnAla 168
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTTGGACATTATACACAGGTTGTTGTTCTACTCTTCATACTTCGTTGGATGTGGAAATGCC 480

Qy     169 TyrCysProAsnGlnLysValLysTyrTyrTyrValCysGlnThrTyrCysProAlaGly 188
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TACTGTCCCAATCAAAAGTTCTAAAATACTACTATGTTTGCCAATATTGTCCCTGCTGGT 540

Qy     189 AsnTrpAlaAsnArgLeuTyrValProTyrGluGlnGlyAlaProCysAlaSerCysPro 208
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AATTGGGCTAATAGACTATATGTCCCTTATGGAACAAGGAGCACCTTGTCCCAGTTGCCCA 600

Qy     209 AspAsnCysAspAspGlyLeuCysThrAsnGlyLysLysTyrGluAspLeuTyrSerAsn 228
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GATAACTGTGACGADGGACTATGCACCAATGGTTGCAAGTACGAAGATCTCTATAGTAAC 660

Qy     229 CysLysSerSerLeuLysSerLeuThrLeuThrCysLysHisGlnLeuValArgAspSerCysLys 248
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TGTAAAAGTTTGAAGCTCACATTAACCTGTAAACATCAGTTGGTCAGGGACAGTTGCAAG 720

Qy     249 AlaSerCysAsnCysSerSerAsnSerIleTyr 258
           |||||||||||||||||||||||||||||||||||
Db     721 GCCTCCTGCAATTGTTCAAACAGCATTTAT 750
```

```
TITLE       The human cysteine-rich secretory protein (CRISP) family. Primary
            structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3
  JOURNAL   Eur. J. Biochem. 236 (3), 827-836 (1996)
  MEDLINE   96270732
   PUBMED   8665901
REFERENCE   2  (bases 1 to 2128)
  AUTHORS   Haendler,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-1996) B. Haendler, Schering AG, ICMB, S109/517,
            13342 Berlin, FRG
FEATURES             Location/Qualifiers
     source          1..2128
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="testis cDNA library"
     gene            16..753
                     /gene="CRISP-3"
     CDS             16..753
                     /gene="CRISP-3"
                     /codon_start=1
                     /product="Cysteine-rich secretory protein-3"
                     /protein_id="CAA64527.1"
                     /db_xref="GI:1262819"
                     /db_xref="SWISS-PROT:P54108"
                     /translation="MTLFPVLLFLVAGLLLPSFPANDKDPAFTALLTTQTQVQREIVN
                     KHNELRRAVSPPARNMLKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYM
                     SSASSWSQAIQSWFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGGCNAYCPNQ
                     KVLKYYYVCQYCPAGNWANRLYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKS
                     LKTLTCKHQLVRDSCKASCNCSNSIY"
     sig_peptide     16..75
                     /gene="CRISP-3"
     polyA_signal    2084..2089
                     /gene="CRISP-3"
BASE COUNT      734 a     397 c     380 g     617 t
ORIGIN
```

```
US-09-698-781-3 (1-258) x HSCRISP3G (1-2128)

Alignment Scores:
Pred. No.:                                  Length:         2128
Score:              3.68e-131               Matches:         249
                    1386.00                 Conservative:      0
Percent Similarity: 99.60%                  Mismatches:        1
Best Local Similarity: 99.60%               Indels:            1
Query Match:        96.52%                  Gaps:              0
DB:
```

```
Qy    9  LeuGluThrThrAlaMetThrLeuPhePheProValLeuLeuPheLeuValAlaGlyLeuLeu  28
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  CTGGAAACCACTGCAATGACAATTATTCCCAGTGCTGTTGTTCCTGTTGTTGCTGGGCTGCTT  60

Qy   29  ProSerPheProAlaAsnGluAspLysAspProAlaPheThrAlaLeuLeuThrThrGln  48
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CCATCTTTTCCAGCAAATGAAGATAAGGATCCCGCTTTTACTGCTTTGTTAACCACCACCAA  120

Qy   49  ThrGlnValGlnArgGluIleValAsnLysHisAsnGluLeuArgArgAlaValSerPro  68
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  ACACAAGTGCAAAGGGAGATTGTGAATAAGCACAATGAACTGAGGAGAGCAGTATCTCCC  180

Qy   69  ProAlaArgAsnMetLeuLysMetGluTrpAsnLysGluAlaAlaAlaAsnAlaGlnLys  88
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CCTGCCAGAAACATGCTGAAGATGGAATGGAACAAAGAGGCTGCAGCAAATGCCCAAAAG  240

Qy   89  TrpAlaAsnGlnCysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLys  108
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TGGGCAAAACAGTGCAATTACAGACACAGTAACCCAAAGGATCGAATGACAAGTCTAAA  300

Qy  109  CysGlyGluAsnLeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSer  128
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  TGTGGTGAGAATCTCTACATGTCAAGTGCCTCACGGTCACAGGTCACAAGCAATCCAAAGC  360

Qy  129  TrpPheAspGluTyrAsnAspPheAspPheGlyValGlyProLysThrProAsnAlaVal  148
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  TGGTTTGATGAGTACAACGATTTTGACTTTGGTGTAGGGCCCAAAGACTCCCAACGCAGTG  420

Qy  149  ValGlyHisTyrThrGlnValValTrpTyrSerSerTyrLeuValGlyGlyCysAsnAla  168
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GTTGGACATTATACACAGGTTGTTTGGTACTCTTCATACCTCGTTGGAGTGTGAAAATGCC  480

Qy  169  TyrCysProAsnGlnLysValLeuLysTyrTyrTyrValCysGlnTyrCysProAlaGly  188
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TACTGTCCCAAATCAAAAAGTTCTAAAATACTATATGTTTGCCAATATTGTTGCTGGT  540

Qy  189  AsnTrpAlaAsnArgLeuTyrValProTyrGluGlnGlyAlaProCysAlaSerCysPro  208
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  AATTGGGCTAATAGACTATATGTCCCTTATGAACAAGGAGCACCTTGTGCCAGTTGCCCA  600

Qy  209  AspAsnCysAspAspGlyLeuCysThrAsnGlyCysLysTyrGluAspLeuTyrSerAsn  228
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  GATAACTGTGACGATGGACTATCCACCAATGGTTGCAAGTATGAAGATCTATATAGTAAC  660

Qy  229  CysLysSerLeuLysThrLeuThrCysLysHisGlnLeuValArgAspSerCysLys  248
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  TGTAAAAGTTTGAAGCTCACATTAACCTGTAAACATCAGTTGGTCAGGGACAGTTGCAAG  720

Qy  249  AlaSerCysAsnCysSerAsnSerIleTyr  258
         |||||||||||||||||||||||||||
Db  721  GCCTCCTGCAATTGTTCAAACAGCATTTAT  750
```

```
RESULT 5
HUMTPX1A
LOCUS       HUMTPX1A     1335 bp    mRNA    linear    PRI   14-JAN-1995
DEFINITION  Human testis-specific protein (Tpx-1) mRNA, complete cds.
ACCESSION   M25532
VERSION     M25532.1  GI:339882
KEYWORDS    testis-specific protein.
SOURCE      Human adult testis, cDNA to mRNA, clones H4-1 and p3-1.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1335)
  AUTHORS   Kasahara,M., Gutknecht,J., Brew,K., Spurr,N. and Goodfellow,P.N.
  TITLE     Cloning and mapping of a testis specific gene with sequence
            similarity to a sperm-coating glycoprotein gene
  JOURNAL   Genomics 5 (3), 527-534 (1989)
  MEDLINE   90129048
   PUBMED   2613236
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by M.Kasahara, 09-JUN-1989.
FEATURES             Location/Qualifiers
     source          1..1335
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="6p21-qter"
     gene            1..1335
                     /gene="GAPDL5"
     mRNA            <1..1335
                     /gene="GAPDL5"
     CDS             190..921
                     /gene="GAPDL5"
                     /product="testis-specific protein mRNA"
                     /note="testis-specific protein precursor"
                     /codon_start=1
                     /protein_id="AAA61220.1"
                     /db_xref="GI:339883"
                     /db_xref="GDB:G00-119-257"
                     /translation="MALLPVLFLFVTVLLPSLPAEGKDPAFTALLTTQLQVQREIVNKH
                     NELRKAVSPPASNMLKMEWSREVTTNAQRWANKCTLQHSDEPEDRKTSTRCGENLYMSS
                     DPTSWSSAIQSWYDEIDLDFVVGVGPKSPNAVVGHYTQLVWYSTYQVCGIAXCPNDS
                     LKYYYVCQYCPAGNMNNKNTPYQOGFTPCAGCPDDCDKGLCTNSCQYQDLLSNCDSLK
                     NFAGCEHELLKEKCKATCLCENKIY"
     sig_peptide     190..249
                     /gene="GAPDL5"
                     /note="testis-specific signal peptide (put.); putative"
     mat_peptide     250..918
                     /gene="GAPDL5"
                     /product="testis-specific protein (put.); putative"
BASE COUNT      428 a     279 c     271 g     355 t     2 others
```

```
ORIGIN          Chromosome 6.

Alignment Scores:
Pred. No.:
Score:                      2.46e-92
Percent Similarity:         1002.00
Best Local Similarity:      80.84%
Query Match:                70.50%
                            69.78%

                Length:       1335
                Matches:       184
                Conservative:   27
                Mismatches:     45
                Indels:          5
                Gaps:            3
DB:
```

US-09-698-781-3 (1-258) x HUMPXIA (1-1335)

```
Qy    1  MetLysGlnIleLeuHisProAlaLeuGluThr-Thr------AlaMetThrLeuPhePr  18
         ::|||| ||||||||||||||| |||||| || ||        |||| || |||| ||
Db  144  ATAAAGTAGATATTTCATCCGTCTCAGAAAAACCAACATTTCCAGCAATGGCTTTACTACC  203

Qy   18  oValLeuLeuPheLeuValAlaGlyLeuLeuProSerPheProAlaAsnGluAspLysAs  38
         |||||||||||||||||||||||||||||||||||||||  |||    ||| |||||
Db  204  GGTG---TTGTTTCTGGTTACTGTGCTGCTTCCATCTTTACCTGCA---GAAGGAAAGGA  257

Qy   38  pProAlaPheThrAlaLeuLeuThrThrGlnThrGlnValGlnArgGluIleValAsnLy  58
         ||||||||||||||||||||||::||||||||||||||||||||||::|||||||||
Db  258  TCCCGCTTTTACTGCTTTGTTAACCACCCAGTTGCAAGTGCAAAGGGAGATTGTAAATAA  317

Qy   58  sHisAsnGluLeuArgArgAlaValSerProProAlaAlaArgAsnMetLeuLysMetGluTr  78
         |||||||||||||::||||||||||||||||||||||||||||||||||||||
Db  318  ACACAATGAACTAAGGAAAGCAGTCTCCACCTGCAGCCAGTAACATGCTAAAGATGGAATG  377

Qy   78  pAsnLysGluLeuAlaLeuAlaAsnAlaGlnLysTrpAlaAsnGlnCysAsnTyrArgHisSe  98
         ||||||||||||||||||||||||||||||||||||||||||||::||||||
Db  378  GAGCAGAGAGGTAACACGAATGCCCAAAGGTGGGCCAAAACAAGTGCACTTTACAACATAG  437

Qy   98  rAsnProLysAspArgMetThrSerLeuLysCysGlyLuAsnLeuTyrMetSerSerAl  118
         ||||||||||||::|||||||||||||||||||||||||::||||||||
Db  438  TGATCCAGGAGGACCCGAAAACCAGTACAAGATGTGGTGAGAATTCTATATGTCAAGTGA  497

Qy  118  aProSerSerTrpSerSerGlnAlaIleGlnSerTrpPheAspGluIleTyrAsnAspPheAspPh  138
         |||||||||||::|||||||||||||||||||||||||||||||||  ::
Db  498  CCCTACTCCTGGTCTTCTCGCAATCCAAAGCTGGTATGACGAGATCCTAGATTTGTCTA  557

Qy  138  eGlyValGlyProLysThrProAsnAlaValValGlyHisTyrThrGlnValValIleTrp  158
         |||||||||||||||||||::||||||||||||||::|||||||||||||
Db  558  TGGTGTAGGACCAAAGAGTCCCAATCCAGTTGTTGGACATTATACTCAGCTTGTTGGTA  617

Qy  158  rSerSerTyrLeuValGlyCysGlyAsnAlaTyrCysGlyProAsnGlnLysValLysTy  178
         ||||||||||||||||::|||||||||::|||||||||||||||||||
Db  618  CTCGACTTACCAGGTAGGGCTGTGGAANTTGCCTACTGTCCCAATCAAGATAGTCTAAATA  677

Qy  178  rTyrTyrValCysGlnTyrCysProAlaGlyAsnTrpAlaAsnArgLeuTyrValProTy  198
         ||||||||||||||||||||||        ||||||
Db  678  CTACTATGTTGCCAATAATGTCTGCTGGTAATAATATGAATAAGAAGAATACCCCGTA  737

Qy  198  rGluGlnLnGlyAlaProCysAlaSerCysProAspAspAsnCysAspAspGlyLeuCysThrAs  218
         ||||||||||||::||||||||||||::|||||||||||||||||||
Db  738  CCAACAAGGAACACCTTGTGCCGGTTCCCTGATGACTGTGACAAGGACTATGCACCAA  797

Qy  218  nGlyCysLysLysTyrGluAspTyrSerAsnCysLysSerLeuLysSerLeuThrThrCy  238
         |||||||||||||||::|||||||||||||::||||||||||||||
Db  798  TAGTTGCCGGTATCAAGATCCTCAGTAACTGTGATTCCTTGAAGATACAGCTGGCTG  857

Qy  238  sLysHisGlnIleLeuValAlaArgAspSerCysLysLysAlaSerCysAsnCysSerAsnSerIleTyr  258
         |::||||||||||::|||||||||||||||||||::||||||
Db  858  TGAACATGAGTTACTCAAGGAAAGTGCAGGCTACTTGCCTATGTGTGAGAACAAAATTTA  917

Qy  258  r  258
         |
Db  918  C  918
```

```
Alignment Scores:
Pred. No.:
Score:                      2.56e-92
Percent Similarity:         1002.00
Best Local Similarity:      80.84%
Query Match:                70.50%
                            69.78%

                Length:       1380
                Matches:       184
                Conservative:   27
                Mismatches:     45
                Indels:          5
                Gaps:            3
DB:
```

US-09-698-781-3 (1-258) x BC022011 (1-1380)

```
Qy    1  MetLysGlnIleLeuHisProAlaLeuGluThr-Thr------AlaMetThrLeuPhePr  18
         ::|||| ||||||||||||||| |||||| || ||        |||| || |||| ||
Db  173  ATAAAGTAGATATTTCATCCGTCTCAGAAAAACCAACATTTCCAGCAATGGCTTTACTACC  232

Qy   18  oValLeuLeuPheLeuValAlaGlyLeuLeuProSerPheProAlaAsnGluAspLysAs  38
         |||||||||||||||||||||||||||||||||||||||  |||    ||| |||||
Db  233  GGTG---TTGTTTCTGGTTACTGTGCTGCTTCCATCTTTACCTGCA---GAAGGAAAGGA  286
```

RESULT 6
LOCUS   BC022011   1380 bp   mRNA   linear   PRI 24-JAN-2002
DEFINITION Homo sapiens, testis specific protein 1 (probe H4-1 p3-1), clone
            MGC:26358 IMAGE:4826427, mRNA, complete cds.

```
COMMENT              Related sequence J04741.
                     13342 Berlin, FRG
         JOURNAL     Submitted (18-JAN-1996) B. Haendler, Schering AG, ICMB, S109/517,
         TITLE       Direct Submission
         AUTHORS     Haendler,B.
   REFERENCE         2  (bases 1 to 1406)
         PUBMED      8665901
         MEDINE      96270732
         JOURNAL     Eur. J. Biochem. 236 (3), 827-836 (1996)
                     structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3
   TITLE             The human cysteine-rich secretory protein (CRISP) family. Primary
                     Donner,P. and Schleuning,W.D.
   AUTHORS           Kratzschmar,J., Haendler,B., Eberspaecher,U., Roosterman,D.,
   REFERENCE         1  (bases 1 to 1406)
                     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   ORGANISM          Homo sapiens
   SOURCE            Homo sapiens.
   KEYWORDS          CRISP-2 gene; cysteine-rich secretory protein-2/type I.
   VERSION           X95539.1  GI:1262816
   ACCESSION         X95539
   DEFINITION        H.sapiens mRNA for cysteine-rich secretory protein-2/type I.
   LOCUS             HSCRISP2I           1406 bp   mRNA   linear    PRI 12-APR-1996
RESULT 7
   HSCRISP2I
```

```
Db  947  C  947
Qy  947  |   258
Db  258  r  258
Qy  258  |

Qy  887  TGAACATGAGTTACTCAAGGAGAAAAGTCGAAGGCTACTGCCTATGTGAGACAAAAATTTA 946
         |:.:||:||||::|::||||||||||||||||||||||||||||| ||| |||||||||
Db  827  TAGTTCCCAGTATCAAGATCTCCTAAGTAACCTGTGATTCCTTGAAGAATACAGCTGGCTG 886

Qy  238  sLysHisGlnLeuValArgAspSerCysLysAlaSerCysAsnCysSerAsnSerIleTy  258
         ||:::|||||||||||:|::||||||||||||||||||||:::|||
Db  767  CCAACAAGGAACACCTTGTCGCCGGTTGCCCTGATGACTGTGACAAAGGACTATGCACCAA 826

Qy  198  rGluGlnGlyAlaProCysAlaSerCysProAsnCysAsnCysAspAspCysGlyLeuCysThrAs 218
         |.:||||||||||||||||||||||||||||||||||||||||
Db  707  CTACTATGTTTGCCAATATTGCCTCGTTGTAATAATATGAATAGAAGAATACCCCGTA 766

Qy  178  rTyrTyrValCysGlnTyrCysProAlaGlyAsnTrpAlaAsnArgLeuTyrValProTy 198
         |||::||:.||::||||||||||||||||||||||||::|||
Db  647  CTCGACTTACCAGGTAGGCTGTGGAAGTTGCCTACTGTCCCAATCAAGATAGTCTAAATA 706

Qy  158  rSerSerTyrLeuValGlyCysGlyAsnAlaThrTyrCysProAsnGlnLysValLeuLysTy 178
         |||.:||||||||||||||||||||||||||||||||||||||
Db  587  TGGTGTAGGACCAAAAGAGTCCCAAATGCAGTTGTTGGACATATTATACTCAGCTTGTTTGGTA 646

Qy  138  eGlyValGlyProLysThrProAsnAlaValAlaValGlyHisTyrThrGlnValValTrpTy 158
         :||||||||||||||||||||||||||||::||||||||||||
Db  527  CCCTACTTCCTGGTCTTCTGCAATCCAAAGCTGGTATGACAGGACATCCTAGATTTGTCTA 586

Qy  118  aProSerSerTrpSerGlnAlaIleGlnSerTrpPheAspGluTyrAsnAspPheAsph 138
         |:::|||||||||||||||||||||||||||||||||
Db  467  TGATCCAGAGGACCCGCAAAAACCAGTACAAGATGTGGTGAGAAATCTCTATATGTCAAGTGA 526

Qy  98   rAsnProLysAspArgMetThrSerLeuLysCysGlyGluAsnLeuTyrMetSerSerAl 118
         |.::||||||||||||||||||::||||||||||||||::||||||
Db  407  GAGCAGAGGAGGTAACAACGAATGCCCAAAGGTGGGGCAAACAAGTGCACTTTACAACTAG 466

Qy  78   pAsnLysGluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnTyrArgHisSe 98
         |.::||||| ||||||     ||||||::::|||
Db  347  ACACAAATGAACTAAGGAAAAGCTCTCCACCTGCCAGTAACATCGTAAGAGATGGAATG 406

Qy  58   sHisAsnGluLeuArgArgAlaValSerProProAlaArgAsnMetLeuLysMetGluTr 78
         |.::||||||||||||||||||||::|||||||||||||||||
Db  287  TCCCGCTTTACTGCTTTGTTAACCACCAGTTGCAAGTGCAAAGGGAGATTGTAAATAA 346

Qy  38   pProAlaPheThrAlaLeuLeuThrThrGlnThrGlnValGlnValLeuAsnLy 58
         |||||||||||||||||||||||||||||||:|||||||||
```

```
FEATURES             Location/Qualifiers
     source          1..1406
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="testis cDNA library"
     gene            241..972
                     /gene="CRISP-2"
     CDS             241..972
                     /gene="CRISP-2"
                     /codon_start=1
                     /product="cysteine-rich secretory protein-2/type I"
                     /protein_id="CAA64526.1"
                     /db_xref="GI:1262817"
                     /db_xref="SWISS-PROT:P16562"
                     /translation="MALLPVLFLVTLLLPSLPAEGKDPAFTALLTLQLVQREIVKH
                     NELRKAVSPPASNMLKMEWSREVTTNAQRWANKCTLQHSDPEDRKTSTRCGENLYMSS
                     DPTSWSSAIQSWYDEILDFVYGVGPKSPNAVGHYTQLVWYSTYQVCGIAYCPNDDS
                     LKYYYCQYCPAGNNMRKNTPYQQGTPCAGCPDDCDKGLCTNSCQYQDLLSNCDSLK
                     NTAGCEHELLKEKCKATCLCENKIY"
     sig_peptide     241..300
                     /gene="CRISP-2"
     polyA_signal    1362..1367
                     /gene="CRISP-2"
ORIGIN
BASE COUNT        458 a        293 c        289 g        366 t

Alignment Scores:
     Pred. No.:                  1
     Score:                      2.62e-92
     Percent Similarity:         1002.00
     Best Local Similarity:      80.84%
     Query Match:                70.50%
                                 69.78%
                                 9

                                          Length:      1406
                                          Matches:     184
                                          Conservative: 27
                                          Mismatches:  45
                                          Indels:      5
                                          Gaps:        3

DB:

US-09-698-781-3 (1-258) x HSCRISP2I (1-1406)
```

```
Qy  1    MetLysGlnIleLeuHisProAlaLeuGluThr-Thr------AlaMetThrLeuPhePr 18
         ||:.:||:    ||||||||||||||||::||| :||   ----||||||||||
Db  195  ATAAAGTAGATATTTCATCCTGCTCAGAAAACAACATTCCACCAATGGCTTTACTACC 254

Qy  18   oValLeuLeuPheLeuValAlaGlyLeuLeuProSerPheProAlaAsnGluAspLysAs 38
         ||||  ||||||||||||||||||||||||||||||||||||||
Db  255  GGTG---TTGTTTCTGGTTACTGTGCTCCTTCCATCTTTACCTGCA---GAAGGAAAGGA 308

Qy  38   pProAlaPheThrAlaLeuLeuThrThrGlnThrGlnValGlnValLeuAsnLy 58
         |||||||||||||||||||||||||||:||||||||||||||||||
Db  309  TCCCGCTTTACTGCTTTGTTAACCACCAGTTGCAAGTGCAAAGGGAGATTGTAAATAA 368

Qy  58   sHisAsnGluLeuArgArgAlaValSerProProAlaArgAsnMetLeuLysMetGluTr 78
         ||||||||||||||||:||||||||||::|||||||||||||||||
Db  369  ACACAAATGAACTAAGGAAAAGCAGTCTCTCCACCTGCCAGTAACATGCTAAAGATGGAATG 428

Qy  78   pAsnLysGluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnTyrArgHisSe 98
         |.::||||:|||      |||::|||::|||
Db  429  GAGCAGAGGAGGTAACAACGAATGCCCAAAGGTGGGGCAAACAAGTGCACTTTACAACATAG 488

Qy  98   rAsnProLysAspArgMetThrSerLeuLysCysGlyGluAsnLeuTyrMetSerSerAl 118
         |.::||||||||||||||||||::||||||||||||::||||||
Db  489  TGATCCAGAGGACCCGCAAAAACCAGTACAAGATGTGGTGAGAAATCTCTATATGTCAAGTGA 548

Qy  118  aProSerSerTrpSerGlnAlaIleGlnSerTrpPheAspGluTyrAsnAspPheAsph 138
         |:::||||||||||||||||||||||||||||||||:
Db  549  CCCTACTTCCTGGTCTTCTGCAATCCAAAGCTGGTATGACAGGACATCCTAGATTTGTCTA 608

Qy  138  eGlyValGlyProLysThrProAsnAlaValAlaValGlyHisTyrThrGlnValValTrpTy 158
         :||||||||||||||||||||||||||||::||||||||||||
Db  609  TGGTGTAGGACCAAAAGAGTCCCAAATGCAGTTGTTGGACATTATACTCAGCTTGTTTGGTA 668

Qy  158  rSerSerTyrLeuValGlyCysGlyAsnAlaThrTyrCysProAsnGlnLysValLeuLysTy 178
         |||.:||||||||||||||||||||||||||||||||||||||
Db  669  CTCGACTTACCAGGTAGGCTGTGGAAGTTGCCTACTGTCCCAATCAAGATAGTCTAAATA 728

Qy  178  rTyrTyrValCysGlnTyrCysProAlaGlyAsnTrpAlaAsnArgLeuTyrValProTy 198
         |||||||||||||||||||||||||||||||||||||
```

```
Db  729 CTACTAGTTTGCCAATATTGTCCTGCTGGTAATAATATGAATAGAAAGAATACCCGTA 788
Qy  198 rGluGlnGlyAlaProCysAlaSerCysProAspAsnCysAspAspGlyLeuCysThrAs 218
        |::||||||      |||||| ||| ||| ||||||||||||||||| |||||||
Db  789 CCAACAAGGAACAACCTTGTGCCCGTTGCCCTGATGACTGTGACAAAGGACTAAGCACCAA 848
Qy  218 nGlyCysLysTyrGluAspLeuTyrSerAsnCysLysSerLeuLysLeuThrLeuThrCy 238
        ||||||||||:|||| ||| ||||| |||||| ||||||||||||||||||||| |||
Db  849 TAGTTGCCAGTATCAAGATCTCCTAAGTAAACGTGACTTCCTTGAAGAATACAGCTGGCTG 908
Qy  238 sLysHisGlnLeuValArgAspSerCysLysAlaSerCysAsnCysSerAsnSerIleTy 258
        |::||:|||:|| || ::::|| |||||| |||||| ||| ||| |||||||||||||
Db  909 TGAAACATGAGTTACTCAAGGAAAAGTGCAAGGCTACTTGCCTATGTGAGACAAAATTTA 968
Qy  258 r 258
        |
Db  969 C 969
```

RESULT 8
CPU35712

```
LOCUS       CPU35712     1388 bp    mRNA    linear   ROD 18-JUN-1996
DEFINITION  Cavia porcellus acrosomal autoantigen 1 mRNA, complete cds.
ACCESSION   U35712
VERSION     U35712.1  GI:1016755
KEYWORDS    .
SOURCE      Cavia porcellus.
  ORGANISM  Cavia porcellus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
REFERENCE   1  (bases 1 to 1388)
  AUTHORS   Foster,J.A. and Gerton,G.L.
  TITLE     Autoantigen 1 of the guinea pig sperm acrosome is the homologue of
            mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich
            secretory protein (CRISP) family
  JOURNAL   Mol. Reprod. Dev. 44 (2), 221-229 (1996)
  MEDLINE   96354287
   PUBMED   9115720
REFERENCE   2  (bases 1 to 1388)
  AUTHORS   Foster,J.A. and Gerton,G.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1995) James A. Foster, Department of Obstetrics
            and Gynecology, Division of Reproductive Biology, University of
            Pennsylvania, 36th and Hamilton Walk, Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1..1388
                     /organism="Cavia porcellus"
                     /strain="Hartley"
                     /db_xref="taxon:10141"
                     /cell_type="sperm acrosome"
                     /tissue_lib="testis"
     CDS             185..919
                     /codon_start=1
                     /product="autoantigen 1"
                     /protein_id="AAC52616.1"
                     /db_xref="GI:1016756"
                     /translation="MALLPVVFLITMLLPCVLTNGKDPAFTALITTQSQVQNEIINK
                     HNQLRKSVTPPASNMLKMEWSREAAVNAQKWANRCTLVHSNPDDRKTSTKCGENLYMS
                     SDPSSWSDAIQSWFDESQDFTFGVGPKSHNAVVGHYTQLVWYSSYLVCGIAYCPNQD
                     SLKYYVCQYCPAGNNVVTKNTPYKQGIPCASCPGHCENGLICTNSCEYEDLLSNCESL
                     KNTAGCEHQLLVEKCKATCEDKIY"
     misc_feature    247..309
                     /note="encodes region previously determined by
                     microsequencing (Hardy et al. Biol Reprod Vol. 38, p.
                     423-437, 1988), PIR Accession Number A60802"
BASE COUNT      444 a    276 c    285 g    383 t
ORIGIN
```

Alignment Scores:

Pred. No.:	2.49e-90	Length: 1388
Score:	982.50	Matches: 179
Percent Similarity:	79.69%	Conservative: 29
Best Local Similarity:	68.58%	Mismatches: 49
Query Match:	68.42%	Indels: 4

DB: 10   Gaps: 2

US-09-698-781-3 (1-258) x CPU35712 (1-1388)

```
Qy    1 MetLysGlnIleLeuHisProAlaLeuGlyThr-Thr------AlaMetThrLeuPhePr 18
        ::|||||||||||::||:|||||          |||        |||| ||| |||||
Db  139 ATAAAGCAGATAGTTCACTCTGGTGTGGAGAACCAACATTCCAGACGAAATGGCCTTACTCC 198
Qy   18 oValLeuLeuPheLeuValAlaGlyLeuLeuProSerPheProAlaAsnGluAspLysAs 38
        :    ||: ::|  ||| |||||||||| ||  |||||||||||| |:||||||||||
Db  199 AGTGGTGGTGTTCTTCATTACTATGCTGCTGCCATGTGTGTACTCACTAAT---GGAAAGGA 255
Qy   38 pProAlaPheThrAlaLeuLeuThrThrGlnThrThrGlnThrArgGluGluIleAsnLy 58
        |||||| ||||||||||||||||||:|| |||    ||| ||||||||:||||||:|||
Db  256 TCCAGCATTTACCCGTTGATAACCACCCAAATCTCAAGGTCCAAAATGAGATTATAAATAA 315
Qy   58 sHisAsnGlnLeuArgArgAlaValSerProProAlaArgAsnMetLeuLysMetGluTr 78
        ||||:|||:|||||| |||||||| |||||||||||||:|||||||||||||||||||
Db  316 ACACAATCAACTGCGGAAAATCTGTACACCCCTTCCTGCCAGCAAATATGCTAAAGACGAATG 375
Qy   78 pAsnLysGluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnGlnCysSerAsnTyrArgHisSe 98
        ||||||||||| ||: |||| |||:|||||||||||:|||||||||| |||:|| ||||
Db  376 GAGCAGAGAGCCAGCAGTAAAATGCACAAAACTGGGCAAATAGGTGTACATTAGTACACAG 435
Qy   98 rAsnProLysAspArgMetThrSerLeuLysCysGlyGluAsnLeuTyrMetSerSerAl 118
        |||||| :||||| ||||||:||||||||||| ||| ||||||||||||||:|||
Db  436 TAATCCAGACGATCGAAAAACCAGCACCAAAATGTGGTGAGAATCTCTATATGTCAAGTGA 495
Qy  118 aProSerSerTrpSerGlnAlaIleGlnSerTrpPheAspGluTyrAsnAspPheAspPh 138
        |||||: |||:|| |||||||| |||||||| |||||||||:||:|| ||||||:|||||
Db  496 CCCTAGCTCCTGGTCACAGATGCAATCCAAAAGCTGGTTTGATGAGGCCAAGAATTTTACCTT 555
Qy  138 eGlyValGlyProLysThrProAlaAsnAlaValValGlyHisTyrThrGlnValValTrpTy 158
        :|||| ||:||||||||||| ||: |||||||| ||| |||:||||||||||||||||
Db  556 TGGAGTAGGACCCAAAACGATCACAATGCGGTTGTTGGACATTACACACAGCTTGTTGTTGGTA 615
Qy  158 rSerSerTyrLeuValGlyCysGlyAsnAlaTyrCysProAsnGlnLysValLeuLysTy 178
        |||||||||||| || |||||: |||||||| ||| |||||||||||||||||||||
Db  616 TTCGTCTTATCTTGTTGGGCGTGTGGGAATTGCCTACGCCCCAATCAAGACAGTCTAAAATA 675
Qy  178 rTyrTyrValCysGlnTyrCysProAlaGlyLeuAsnAsnArgLeuTyrValProTy 198
        |:|| :||||||||| |||||||| |||||| |||:|| ||||||:|||||
Db  676 CTACTACGTTTGCCAATACTGTCCTGCTGGTAATAATGGTTATACAAAGAACACTCCTTA 735
Qy  198 rGluGlnGlyAlaProCysAlaSerCysProAspAsnCysAspAspGlyLeuCysThrAs 218
        ::||||||     |||||| ||| ||| |||||||||||||||||| |||||||
Db  736 TAAACAAAGGAATACCTTGTGCCAGCTGCCGACATTGTGAAAATGGACATTGCACCAA 795
Qy  218 nGlyCysLysTyrGluAspLeuTyrSerAsnCysLysSerLeuLysLeuThrLeuThrCy 238
        ||||||||||:||| |:: ||| |||||| ||||||||||||||||||||| |||
Db  796 CAGTTGTCAGTATGAAGATCTCCTAGTAACTGCGAGTCTTGAAGAATACAGCTGGCTG 855
Qy  238 sLysHisGlnLeuValArgAspSerCysLysAlaSerCysAsnCysSerAsnSerIleTy 258
        |::||:|||:|| ||: :|| |||||| |||::| ||| ||| |||||||||||||
Db  856 TGAACACATCAATTGCTGTGGAAAGTGCAAGGCTACTTGCAGATGTGAAGATAAAATATA 915
Qy  258 r 258
        |
Db  916 C 916
```

RESULT 9
ECCRISP3

```
LOCUS       ECCRISP3     1295 bp    mRNA    linear   MAM 03-FEB-1998
DEFINITION  Equus caballus mRNA for cysteine-rich secretory protein-3.
ACCESSION   AJ001400
VERSION     AJ001400.1  GI:2388784
KEYWORDS    CRISP-3; cysteine-rich secretory protein-3.
SOURCE      horse.
  ORGANISM  Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 1295)
  AUTHORS   Schambony,A.
```

TITLE   Direct Submission
JOURNAL   Submitted (03-SEP-1997) Schambony A., Institut fuer
            Reproduktionsmedizin, Tieraerztliche Hochschule Hannover, Buenteweg
            15, D-30559 Hannover, GERMANY
REFERENCE   2 (bases 1 to 1295)
AUTHORS   Magdaleno,L., Gasset,M., Varea,J., Schambony,A.M., Urbanke,C.,
            Raida,M., Topfer-Petersen,E. and Calvete,J.J.
TITLE   Biochemical and conformational characterisation of HSP-3, a
            stallion seminal plasma protein of the cysteine-rich secretory
            protein (CRISP) family
JOURNAL   FEBS Lett. 420 (2-3), 179-185 (1997)
MEDLINE   98119394
PUBMED   9459306
FEATURES   Location/Qualifiers
     source   1..1295
                     /organism="Equus caballus"
                     /db_xref="taxon:9796"
                     /sex="Male"
                     /tissue_type="ampulla"
                     /dev_stage="adult"
     gene   59..796
                     /gene="crisp-3"
     CDS   59..796
                     /gene="crisp-3"
                     /codon_start=1
                     /product="cysteine-rich secretory protein-3"
                     /protein_id="CAA04729.1"
                     /db_xref="GI:2388785"
                     /db_xref="SWISS-PROT:O19010"
                     /translation="MALLPVLLFLAAVLLPFFPASGQDPGFAALSITKSEVQKEIVNK
                     HNDLRRTVSPLASNMLKMQWDSKTATNAQWANKCLLQHSKAEDRAVGTMKCGENLFM
                     SSIPNSWSDAIQNWHDEVHDFKYGVGPKTPNAVVGHYTQVVWYSSYRVGCGIAYCPKQ
                     GTLKYYVCQYCPAGNVYNKINTPYEQGTPCARCPGNCDNGLCTNSCEYEDLVSNCDS
                     LKKIAGCEHELLKENCKYTCQCENKIY"
     sig_peptide   59..124
                     /gene="crisp-3"
     polyA_signal   1262..1267
                     /gene="crisp-3"
BASE COUNT   428 a   257 c   245 g   365 t
ORIGIN

US-09-698-781-3 (1-258) x ECCRISP3 (1-1295)

Alignment Scores:
Pred. No.:
Score: 2.58e-87
Score: 952.50
Percent Similarity: 78.76%
Best Local Similarity: 65.25%
Query Match: 66.33%
DB: 4

                                                               Length: 1295
                                                               Matches: 169
                                                               Conservative: 35
                                                               Mismatches: 52
                                                               Indels: 3
                                                               Gaps: 3

Qy   1 MetLysGlnIleLeuHisProAlaLeuGluThrThrAlaMetThrLeuPheProValLeu 20
           ||||||||::|||||||||:||||||||||||||||||| |||||||
Db   1 MetLysGlnIleLeuHisProAlaLeuGluThrThrAlaMetThrLeuPheProValLeu 20

Qy   21 LeuPheLeuValAlaGlyLeuLeuProSerProAlaAsnGluAspLysAspProAla 40
         |||||||||| ||||||||| ||||||||||||| |||||||||
Db   23 GTAAAGCAGATGCTTCACCTTGCTCTGAGACACCAAC---ATGGCATRATTACCAGTGTTG 79

Qy   41 PheThrAlaLeuLeuThrThrGlnThrGlnValGlnIleValAlaAsn 60
         |||||| ||||||||||||||||||||||||||||||||||
Db   80 CTGTTTCTGGGCTCCGGTGTTGCTTCCAATTCTTTCCTGCGAAGT---GGACAGGATCCAGGT 136

Qy   61 GluLeuArgArgAlaValSerProProAlaArgAspAspLysAspProAla 80
         |||| |||||||||||||||||||||||||| |||||||||
Db   137 TTTGCTCGCTTTGTCAATCACCAAAAGCCAAGTCCAAAAGAGATTGTAAATAAACACAAT 196

Qy   81 GluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnGlnCysAsnTyrArgHisSerAspAsn 100
         ||| |||||||||||| |||||||||||||||||||||||||||
Db   197 GACCTAAGGAGAACAAGTCTCTCCACTTGCCACATGCTAAAGACAAGTGCAAATGGACAGC 256

Qy   101 LysAspArg--MetThrSerLeuLysCysGlyGluAsnLeuTyrMetSerSerAlaPro 119
         ||: |||| |||||||||||||||||||||||||||||||||||
Db   257 AAGACACAACAAAATGCCCAAAACTGGGCCAACAAGTGCCTTCTCCAACACAGTAAAGCA 316

Qy   120 SerSerTrpSerGlnAlaIleGlnSerTrpHisAspThrAlaAspGluTyrAsnAspPheGly 139
         ::||||||||| ||||||||||||| |:|:|||||||||||
Db   377 AATTCCTTGGTCAGAATGCCAATCCAAATCAAAAATTTGGGACAGATGAGGTCCATGATTTAAATACGGC 436

Qy   140 ValGlyProLysThrProAsnAlaValValGlyHisTyrThrGlnValValTrpTyrSer 159
         |||||||||||||||||||| ||||||||||||||||||
Db   437 GTGGGCCCAAAGACTCCCAATGCAGTAGTTGGACATTATCACCAGGTTGTTTGGTACTCA 496

Qy   160 SerTyrLeuValGlyCysGlyAsnAlaTyrCysProAsnGlnLysValLeuLysTyrTyr 179
         |||||| ||||||||||||||||||||||||||||||||||||
Db   497 TCTTACCGTGTTGGATGTGGAAATGCGTATTGTCCCAACCAAAGAAGGTCTAAAATACTAC 556

Qy   180 TyrValCysGlnTyrCysProAlaGlyAsnArgLeuTyrValProTyrGlu 199
         ||||||||||| ||||||||||||||||||||| ||||||||||||
Db   557 TATGTTTGCCAAATATTGCCCTGCTGGTAATATATGTGAATAAATACTCCTTATGAA 616

Qy   200 GlnGlyAlaProCysAlaSerCysProAspAspAsnCysAspAspGlyLeuCysThrAsnGly 219
         |||||||||||||||||||||||||| :||||||||||||||||||
Db   617 CAAGGAACACCTTGTGCCCGTGCCCCTGGTAACTGTGACAATGGTCTTTGTACTATCCACAATAGC 676

Qy   220 CysLysTyrGluAspLeuTyrSerAsnCysLysSerLeuLeuThrLeuThrCysLys 239
         |||||||||||||||||||||||||||||||| |||:|
Db   677 TGCGAGTATGAAGATCTCGTTAGTAACTGTGATTCCTTGAAGAAAAATAGCTGGTCTGTGAA 736

Qy   240 HisGlnLeuValArgAspSerCysLysAsnCysSerAsnSerIleTyr 258
         |||||| |||:|||||||||||||| ||| |||||||||
Db   737 CATGAATTGCTCAAGGAAAAACTGCAAGACTACTTGTCAATGTGAAAACAAAATTAC 793

RESULT 10
LOCUS   AB009662   1280 bp   mRNA   linear   ROD 01-AUG-1998
DEFINITION Rattus norvegicus mRNA for testis specific protein, complete cds.
ACCESSION   AB009662
VERSION   AB009662.1 GI:3374579
KEYWORDS   testis specific protein.
SOURCE   Rattus norvegicus
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 1280)
AUTHORS   Maeda,T., Sakashita,M., Ohba,Y. and Nakanishi,Y.
TITLE   Molecular cloning of the rat Tpx-1 responsible for the interaction
            between spermatogenic and Sertoli cells
JOURNAL   Biochem. Biophys. Res. Commun. 248 (1), 140-146 (1998)
MEDLINE   98340864
REFERENCE   2 (bases 1 to 1280)
AUTHORS   Nakanishi,Y.
TITLE   Direct Submission
JOURNAL   Submitted (11-DEC-1997) Yoshinobu Nakanishi, Kanazawa University,
            Faculty of Pharmaceutical Sciences; 13-1 Takara-machi, Kanazawa,
            Ishikawa 920, Japan (tel:076-234-4480, Fax:076-234-4480)
FEATURES   Location/Qualifiers
     source   1..1280
                     /organism="Rattus norvegicus"
                     /strain="Donryu"
                     /db_xref="taxon:10116"
                     /clone="Tpx-1"
                     /sex="male"
                     /cell_type="Spermatogenic cells"
                     /tissue_type="Testis"
                     /dev_stage="20 days after birth"
     CDS   78..809
                     /codon_start=1
                     /product="testis specific protein"
                     /protein_id="BAA32029.1"
                     /db_xref="GI:3374580"
                     /translation="MAWFQVMLFVFAVLLPLPPTEGKDPDFAFLLTTNQIQVQREIIAK
                     HNELRRQVSPPGSNILKMEMNVQQAANARQKWANNCILEHSSTEDRKINIKCGENLYMS
                     TDFTSWRIVIQSWYEEENENFVGVGAKPNSAVGHYTQLVWYSSFKVGCGVAYCPNQDT
                     LKYFYVCHYCPMGNNVNVKKSTPYHQGTPCASCPNNCDNGLCTNSCDFEDLSNCESLK

```
BASE COUNT      416 a    263 c    259 g    342 t
ORIGIN
              polyA_site    1280
                            /note="20 a nucleotides"
              SSAGCKHELLKAKCEATCLCEDKIH"

Alignment Scores:
Pred. No.:              1.68e-77
Score:                  856.00
Percent Similarity:     74.39%
Best Local Similarity:  61.79%
Query Match:            59.61%
DB:                     10

                        Length:        1280
                        Matches:       152
                        Conservative:  31
                        Mismatches:    61
                        Indels:        2
                        Gaps:          2

US-09-698-781-3 (1-258) x AB009662 (1-1280)

Qy    13 AlaMetThrLeuPheProValLeuLeuPheLeuValAlaGlyLeuLeuProSerPhePro 32
         ||||||    ||    ||||||||||||::::  |||||||  |||||||||||
Db    75 GCCATGGCTTGGTTCCAGGTGATGCTGTTTGTCTTTGCTGTGGCCTGCTACCA---TTGCCA 131

Qy    33 AlaAsnGluAspLysAspProAlaPheThrAlaLeuLeuThrThrGlnThrGlnValGln 52
         |||||||||||||||   |||  |||||||||
Db   132 CCCACAGAAGGAAAGGATCCAGACTTCGCTACTTTGACAACCAACCAAATACAAGTTCAA 191

Qy    53 ArgGluIleValAsnLysHisAsnGluLeuArgArgAlaValSerProProAlaArgAsn 72
         |||||||||   |||:::   |||||||||  |||||||||
Db   192 AGAGAGATCATAGCTAAACACAATGAACCTAGCCCCCCTGGCAGCCAC 251

Qy    73 MetLeuLysMetGluTrpAsnLysGluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnGln 92
         ||||||||||||||||||||:::  ||||||||||||
Db   252 ATACTAAAAATGGAATGGAACGTACAAGCACCAGCAAATGCTCAAAAGTGGGCTAATAAC 311

Qy    93 CysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLysCysGlyGluAsn 112
         ||||||||||||||::::||||||||||||
Db   312 TGTAATTTTAGAACACAGTAGTACAGAAGACCGAGAAAATCAATAATAAAATGTGGCGAGAAT 371

Qy   113 LeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSerTrpPheAspGlu 132
         ||||||||||||||||::::|||||||||||||
Db   372 CTCTATATGTCGACTGACCCTACACCCTGGAGAACCGTAATTCAAAGCTGGTTATGAAGAA 431

Qy   133 TyrAsnAspPheAspPheGlyValGlyProLysThrProAsnAlaValValGlyHisTyr 152
         ||:::||||||   |||   ||||||||||
Db   432 AATGAAAACTTCGTTTTCGGCGGTGTAGGAGCTAAA---CCCAATTCCGCTGTCTCGGACACTAC 488

Qy   153 ThrGlnValValTrpTyrSerSerTyrLeuValGlyCysGlyAsnAlaThrCysProAsn 172
         |||||||||||||||::::||||||||||||
Db   489 ACTCAGGTTGTTTGGTATTCAAGTTGGATGTGGAGTTGCTTACNTGCTCCAAT 548

Qy   173 GlnLysValLeuLysTyrTyrTyrTyrValCysGlnTyrCysProAlaGlyAsnTrpAlaAsn 192
         ||    ::|||||||||||||||||||| |||||||
Db   549 CAAGATACCCTGAAAATACTTCTATGTTTGCCATTACTGTCCTATGGGTAACAACGTGATG 608

Qy   193 ArgLeuTyrValProTyrGlyLeuGlnGlyAlaProCysAlaSerCysProAspAsnCysAsp 212
         :::|||  ||||||||||
Db   609 AAAAACAGTACCCCATATCATCAAAGGACACCTTGTCCTAGTTGTCCCAATAACTGTGAT 668

Qy   213 AspGlyLeuCysThrAsnGlyCysLysTyrGluAspLeuTyrSerAsnCysLysSerLeu 232
         ::::|||||||  ||||||||||||
Db   669 AATGGATTGTGCACCAATAGCTGTAAATATGAAGATTTACTTAGTAACTGTAAATCCTTG 728

Qy   233 LysLeuThrLeuThrCysLysHisGlnLeuValAlaArgAspSerCysLysSerAlaSerCysAsn 252
         ||   ::|||||||||||||||:::|||||||
Db   729 AAGAGTTCAGCAGGCTGTAAACATGAGTTGCTCAGGCAAAGTGTGAGGCTACTTGCCTA 788

Qy   253 CysSerSerAsnSerIleTyr 258
         ||    :::  ||||||
Db   789 TGTGAAGACAAAAATTCAT 806
```

```
RESULT 11
AF078552
LOCUS       AF078552    1432 bp    mRNA    linear    ROD 11-AUG-1999
DEFINITION  Rattus norvegicus cysteine-rich secretory protein-2 CRISP-2 (tpx-1)
            mRNA, complete cds.
```

```
ACCESSION   AF078552
VERSION     AF078552.1  GI:5725525
KEYWORDS    .
SOURCE      Rattus norvegicus.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1432)
AUTHORS     O'Bryan,M.K., Loveland,K.L., Herszfeld,D., McFarlane,J.R.,
            Hearn,M.T. and de Kretser,D.M.
TITLE       Identification of a rat testis-specific gene encoding a potential
            rat outer dense fibre protein
JOURNAL     Mol. Reprod. Dev. 50 (3), 313-322 (1998)
MEDLINE     98284327
PUBMED      9621307
REFERENCE   2  (bases 1 to 1432)
AUTHORS     O'Bryan,M.K., Loveland,K.L., Herszfeld,D., McFarlane,J.R.,
            Hearn,M.T. and de Kretser,D.M.
TITLE       Direct Submission
JOURNAL     Submitted (15-JUL-1998) Institute of Reproduction & Development,
            Monash University, Monash Medical Centre, 246 Clayton Rd., Clayton,
            Victoria 3168, Australia
FEATURES             Location/Qualifiers
     source          1..1432
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     gene            1..1432
                     /gene="tpx-1"
                     /note="similar to mouse and human tpx-1"
     CDS             244..975
                     /gene="tpx-1"
                     /note="putative component of sperm tail outer dense
                     fibers; TPX-1"
                     /codon_start=1
                     /product="cysteine-rich secretory protein-2 CRISP-2"
                     /protein_id="AAD48090.1"
                     /db_xref="GI:5725526"
                     /translation="MAWFQVMLFVFGVLLPLPTEGKDPDFARLTTNQIQVQREIIAK
                     HNLRRQVSPPGSNILKEMVNQAAANAQKWANNCILEHSSTEDRKINIKCGENLYMS
                     TDPTSWTFVIQSWYEEENFVGVGAKPNSAVGHYTLQLVWYSSFKVCGVAYCPNQDT
                     LKYFYVCHYCPMGNNVMKKSTPYHQGTPCASCPNNCDNGLCTNSCDFEDLLSNCESLK
                     SSARCKHELLKAKCEATCLCEDKIH"
BASE COUNT   458 a   301 c   299 g   373 t   1 others
ORIGIN
```

```
US-09-698-781-3 (1-258) x AF078552 (1-1432)

Alignment Scores:
Pred. No.:              3.9e-77
Score:                  853.00
Percent Similarity:     73.98%
Best Local Similarity:  61.38%
Query Match:            59.40%
DB:                     10

                        Length:        1432
                        Matches:       151
                        Conservative:  31
                        Mismatches:    62
                        Indels:        2
                        Gaps:          2

Qy    13 AlaMetThrLeuPheProValLeuLeuPheLeuValAlaGlyLeuLeuProSerPhePro 32
         ||||||    ||    ||||||||||||::::  |||||||  |||||||||||
Db   241 GCCATGGCTTGGTTCCAGGTGATGCTGTTTGTCTTTGCTGTGGCCTGCTACCA---TTGCCA 297

Qy    33 AlaAsnGluAspLysAspProAlaPheThrAlaLeuLeuThrThrGlnThrGlnValGln 52
         |||||||||||||||   |||  |||||||||
Db   298 CCCACAGAAGGAAAGGATCCAGACTTCGCTACTTTGACAACCAACCAAATACAAGTTCAA 357

Qy    53 ArgGluIleValAsnLysHisAsnGluLeuArgArgAlaValSerProProAlaArgAsn 72
         |||||||||   |||:::   |||||||||  |||||||||
Db   358 AGAGAGATCATAGCTAAACACAATGAACCTAGCCCCCCTGGCAGCCAC 417

Qy    73 MetLeuLysMetGluTrpAsnLysGluAlaAlaAlaAsnAlaGlnLysTrpAlaAsnGln 92
         ||||||||||||||||||||:::  ||||||||||||
Db   418 ATACTAAAAATGGAATGGAACGTACAAGCACCAGCAAATGCTCAAAAGTGGGCTAATAAC 477
```

Left column — alignment (query protein vs. MUSTPX1A nucleotide)

```
Qy   93 CysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLysCysGlyGluAsn 112
           ||||||           ::|||||||||||||||||
Db  478 TGTATTTTAGAACACAGTAGTACAGAAGACCGAGAAAATCAATATAAAAATGTGGCAGAAT 537

Qy  113 LeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSerTrpPheAspGlu 132
           :||||| ||||::||
Db  538 CTCTATATGTCGACTAGCCCTACATCCTGGAGAACCGTAATTCAAAGCTGGTTTGATGAAGAA 597

Qy  133 TyrAsnAspPheAspPheGlyValGlyProLysThrProAsnAlaValValGlyHisTyr 152
           ||||||||||||||
Db  598 AATGAAAAACTTCGTTTCGGGGTAGGAGCTAAA---CCCAATTCCGCTGTCGGACACTAC 654

Qy  153 ThrGlnValValTrpTyrSerSerTyrLeuValGlyCysGlyAsnAlaTyrCysProAsn 172
           |||||||||||||||||||||||
Db  655 ACTCAGGTTGTTTGGTATTCATCTTACCTTGTTGGTTGCGGAAATGCTTACTGTCCCAAT 714

Qy  173 GlnLysValLeuLysTyrTyrTyrTyrValCysGlnTyrCysProAlaGlyAsnTrpAlaAsn 192
           ||| ||||:|||||||||
Db  715 CAAGATACCCCTGAAATACTTCTAYGTTTGCCAGTGTTGCCCAGCTGGTAACAACGTGATG 774

Qy  193 ArgLeuTyrValProTyrGluGlnGlyAlaProLysAlaSerCysProAlaSerProAspAsnCysAsp 212
           ||:||||||||
Db  775 AAAAAGACTACCCCATACATCAACAAGGGACACCTTGTGCTTGTCCCAGTAGTTAACTAACTGTGAT 834

Qy  213 AspGlyLeuCysThrAsnGlyCysLysTyrArgCysSerProAlaGlyAsnTrpAlaAsn 232
           |||:||||||||||||
Db  835 AATGGATGTTGCACCAATAGCTGTAAGTTTAGAAGATTTACTAGTAACGTGAATCCTTG 894

Qy  233 LysLeuThrLeuThrCysLysHisGlnLeuValAlaArgAspSerCysLysAlaSerCysAsn 252
           :||:||||||||||||||::|||
Db  895 AAGAGTTCAGCACGCCTGTAAACAATGAGTTGGTCGCAAGAGTCTGAGGCTACTTGCCTA 954

Qy  253 CysSerSerAsnSerIleTyr 258
           ::||          |||:::
Db  955 TGTGAGACAAAATTCAT 972
```

RESULT 12

```
LOCUS       MUSTPX1A              1418 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse testis-specific protein (Tpx-1) mRNA, complete cds.
ACCESSION   M25533 J04741
VERSION     M25533.1  GI:202126
KEYWORDS    testis-specific protein.
SOURCE      Mouse (strain BALB/c) testis, cDNA to mRNA, clone B4(1)4.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1418)
AUTHORS     Kasahara,M., Gutknecht,J., Brew,K., Spurr,N. and Goodfellow,P.N.
TITLE       Cloning and mapping of a testis-specific gene with sequence
            similarity to a sperm-coating glycoprotein gene
JOURNAL     Genomics 5 (3), 527-534 (1989)
MEDLINE     90129048
PUBMED      2612236
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by M.Kasahara, 09-JUN-1989.
FEATURES             Location/Qualifiers
     source          1..1418
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     mRNA            <1..1418
                     /product="testis-specific protein mRNA"
     CDS             214..945
                     /note="testis-specific protein precursor"
                     /codon_start=1
                     /product="testis-specific protein"
                     /protein_id="AAA40472.1"
                     /db_xref="GI:202127"
                     /translation="MAWFQVMLFVFALLLRSPLTEGKDPDPTSLLFDNLQLVQREIVNK
                     HNELRRSVNFPGSDILKMEWSIQATTNAQKNAVGHYTQLVWYSSFKIGCGIAYCPNQDN
                     TDPTLMSTVIQSWYNENEDFYVGVGAKPNSAVGHTQLVWYSSFKIGCGIAYCPNQDN
                     LKYFVCHYCPMGNNVNMKKSTPYQQGTPCASCPNNCENGLCTNSCDFPEDLLSNCESLK
                     TSAGCKHELKTKCQATCLCEDKIH"
     sig_peptide     214..279
                     /note="testis-specific signal peptide (put.); putative"
     mat_peptide     280..942
                     /product="testis-specific protein (put.); putative"
BASE COUNT      460 a    287 c    283 g    388 t
ORIGIN          Chromosome 17.
```

Right column — Alignment Scores

```
Alignment Scores:
Pred. NO.:
Score:                    2.61e-75        Length:        1418
                            835.00        Matches:        147
Percent Similarity:        75.718         Conservative:    40
Best Local Similarity:     59.518         Mismatches:      58
Query Match:               58.158         Indels:           2
DB:                            10         Gaps:             2
```

US-09-698-781-3 (1-258) x MUSTPX1A (1-1418)

```
Qy   12 ThrAlaMetThrLeuPheProValLeuLeuPheLeuValAlaLeuGlyLeuLeuProSerPhe 31
           :|||          ||:::
Db  208 TCGGCCATGGCCTTGGTTCCAACTGATGCTGTCTTTGCTTTGCGTTGCGTGCGATCA--- 264

Qy   32 ProAlaAsnGluAspLysAspLysAspProAlaPheThrAlaLeuLeuThrThrGlnVal 51
           |||:||||||||||||||
Db  265 CCACTCACAGAAGGAAAGGATCCAGACTTTACTCTGTTGTTAACCAACAAACTACAAGTT 324

Qy   52 GlnArgGluIleValAsnLysHisAsnGluLeuArgArgAlaValAlaValSerProAlaArg 71
           |||||||||||||||||:|:|||
Db  325 CAAAGAGAGATCGTAAATAAACACAATGAGCTGAGGAGGATCAGTTAACCCTACTGGCAC 384

Qy   72 AsnMetLeuLysMetGluTrpIleProAsnLysGlnAlaAlaAlaAsnAlaGlnLysTrpAlaAsn 91
           |||||:|||
Db  385 GACATACTAAAAAATGGAATGGAGACTACAACAAACAAATGCTCAAAAAGTGGGCTAAT 444

Qy   92 GlnCysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLysCysGlyGlu 111
           :::|||          |||:::
Db  445 AAGGTATTTTAGACACAGTTCTAAAGATGACCGGAAGATCAATATAAGAGTGTGGAGGTGAG 504

Qy  112 AsnLeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSerTrpPheAsp 131
           :||:|||||||||||||||||||:||:
Db  505 AATCTCTATATGTCAACTGATCCTACCCCTGGTCACCCTCTGGTCCAACTGTAATTCAAAGCTGGTACAAT 564

Qy  132 GluTyrAsnAspPheAspPheGlyValGlyProLysThrProAsnAlaValValGlyHis 151
           |||||||||||||
Db  565 GAAAAATGAAGATTTTGATTTGTTTATGGTGTAGGAGCTAAA---CCCAATTCAGCTGTCGGACAC 621

Qy  152 TyrThrGlnValValTrpTyrSerSerTyrLeuValGlyCysGlyAsnAlaTyrCysPro 171
           |||||||||||||||||||||||:|
Db  622 TACACTCAGGTTGTTTGGTATTCAATCTTTACAAGTTGGTTGCGGAAATGCTTACTGTCCC 681

Qy  172 AsnGlnLysValLeuLysTyrTyrTyrTyrValCysGlnTyrCysProAlaGlyAsnTrpAlaAla 191
           |||:||||||
Db  682 AATCAAGACAACCTGAAAATATTTCTATGTTTGCCATTACTGTCCTATGGGTAACAATGTG 741

Qy  192 AsnArgLeuTyrValProTyrGluGlnGlyAlaProLysAlaSerCysProAlaSerProAspAsnCys 211
           :||:||||||||
Db  742 ATGAAAAGAAAGGCACCCCCATATCAACAAGGAACACCTTGTGCTAGTTGTCCAATAACTGT 801

Qy  212 AspAspAspGlyLeuCysThrAsnGlyCysLysTyrGluAspLeuTyrSerAsnCysLysSer 231
           :|::||||||||||
Db  802 GAAAAATGGGCTGTGCACCAATAGCTGTAAGTTTGAAGATTTACTTAGTAACTGTAAATCC 861

Qy  232 LeuLysLeuThrLeuThrCysLysHisGlnLeuValAlaArgAspSerCysLysAlaSerCys 251
           |||:||||||||||||::|||
Db  862 TTGAAGACTTCACCAGGCTGTAAACATGAGTTGCTCAGAGACAAAAGTGCCAGGCTACTTGT 921

Qy  252 AsnCysSerSerAsnSerIleTyr 258
           :::||          |||:::
Db  922 TTGTGTGAAGACAAAATTCAT 942
```

RESULT 13

```
LOCUS       MUSAEGIA              1380 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse acidic epididymal glycoprotein (Aeg-1) mRNA, complete cds.
ACCESSION   M92849
```

```
VERSION     M92849.1 GI:191737
KEYWORDS    sperm coating glycoprotein; sperm maturation.
SOURCE      Mus musculus (strain C57BL/6J, sub_species domesticus) (library:
            Uni-ZAP) submandibular gland cDNA to mRNA.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1380)
AUTHORS     Vazquez,M., Mizuki,N., Flajnik,M.F., McKinney,E.C. and Kasahara,M.
TITLE       Nucleotide sequence of a nurse shark immunoglobulin heavy chain
            cDNA clone
JOURNAL     Mol. Immunol. 29 (9), 1157-1158 (1992)
MEDLINE     92357056
PUBMED      1495502
REFERENCE   2  (sites)
AUTHORS     Mizuki,N. and Kasahara,M.
TITLE       Mouse submandibular glands express an androgen-regulated transcript
            encoding an acidic epididymal glycoprotein-like molecule
JOURNAL     Mol. Cell. Endocrinol. 89 (1-2), 25-32 (1992)
MEDLINE     93246016
PUBMED      1301383
FEATURES             Location/Qualifiers
     source          1..1380
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /tissue_type="submandibular gland"
                     /tissue_lib="Uni-ZAP"
     gene            1..1380
                     /gene="Aeg-1"
     CDS             22..756
                     /gene="Aeg-1"
                     /gene="Aeg-1"
                     /codon_start=1
                     /product="acidic epididymal glycoprotein"
                     /protein_id="AAA37185.1"
                     /db_xref="GI:191738"
                     /translation="MALMLVLFLAAVLPPSLLQDSSQENRLEKLSTTKMSVQEEIVS
                     SSYLASWSAAIQGWYNEYKDLTDYDVGPGKQPDSVGHYTQVVWNSFQVACGVAECPKN
                     KKMLSCEHELLKKGCKATCLCEGKIH"
     sig_peptide     22..78
                     /gene="Aeg-1"
     polyA_signal    1357..1362
                     /gene="Aeg-1"
BASE COUNT      434 a    263 c    240 g    443 t
ORIGIN
```

```
US-09-698-781-3 (1-258) x MUSAEG1A (1-1380)

Alignment Scores:
Pred. No.:        5.79e-69
Score:            772.50        Length:         1380
Percent Similarity: 70.68%      Matches:        136
Best Local Similarity: 54.62%   Conservative:   40
Query Match:      53.80%        Mismatches:     72
DB:               10            Indels:         1
                                Gaps:           1
```

```
Qy     10 GluThrThrAlaMetThrLeuPheProValLeuPheLeuValAlaGlyLeuLeuPro 29
          |||     |||  ||||    ||||   |||||||||||||||
Db     10 GAAGCCAGCACCATGGCATTAATGCTTGTCGTTCTTCTTTCTGGCTGCTGCTACTGCCCCA 69

Qy     30 SerPheProAlaAsnGluAspLysAsppProAlaPheThrAlaLeuLeuThrThrGlnThr 49
          :::        :::           |||  ||||||::::
Db     70 TCCCTTCTTCAAGATAGCTCTCAGGAAAACGTCTTGAGAAACTTTCAACCACTAAAATG 129

Qy     50 GlnValGlnArgGlnIleValAsnLysHisAsnGluLeuArgAsnLeuSerProPro 69
          ||||||     |||||||||||::::|||||||||::::|||||||||
Db    130 TCAGTCCAAGAAGAGATTGTAAGCAACGCACAACCAATTGAGCAGTAAGGTTTCCTCARCT 189

Qy     70 AlaArgAsnMetLeuLysMetGluTrpAsnLysGluAlaIleAlaLeuAsnAlaGlnLysTrp 89
          :::::|||||||||||||||||||||||   ::|||
```

```
Db    190 GGCAGTGGACTTACTAAAAATGGAATGGAACTATGATGCTCAAGTGATGTCAGCAATGG 249
Qy     90 AlaAsnGlnCysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLysCys
          |||::::|||||||||||||||||||::||:|||::|::|:::|||:||:|::|||::||
Db    250 GCAGACAAGGTGACATTCCAGTCACAGTCCTATAGAACTCAGGACACACTAATTTAAGATGT 309
Qy    110 GlyGluAsnLeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSerTrp 129
          |||||::||||::||::||||||||||||::||:::::|::|||||:::|||::::|||
Db    310 GGGGAGAATTTGTTCATGTCATCTGCACCTTCGTCATGCACATGGTCTTCTGCAATCCAAGG 369
Qy    130 PheAspGluTyrAsnAsppPheAspGlyValGlyProLysThrProAsnAlaLaValVal 149
          ::::||::||:::|||||::::||::|:::|:::::|||::||:::||:||:::::|||::|
Db    370 TATAATGAATACAAAGATCTTAGATATGATGTTGGGCCAAAAGCAACCTGATAGTGGTT 429
Qy    150 GlyHisTyrThrGlnValValIleProTyrSerSerTyrLeuValGlyCysGlyAsnAlaTyr 169
          |||::|::|:|||||::|:||||::||||||||::||||:||:|:|:||||||||
Db    430 GGACATTATACTCAGGTTGTTGTTGGACTACACTGCAACTTTCCAAGTTGCCAGTTGCCTGAA 489
Qy    170 CysProAsnGlnLysValLeuLysTyrTyrTyrValCysGlnTyrCysProAlaGlyAsn 189
          |||||||:::||:::||||||||||||||:::||:||||||||||::||||||||||||
Db    490 TGCCCT---AAAAATCCACTGAGGATACTATTATGTTTGTCACTATTGTCCTGTGGCAAT 546
Qy    210 AsnCysAspAspGlyLeuCysThrAsnGlyCysLysTyrGluAspLeuThrTyrSerAsnCys 229
          ||  ||::||||::|||||||||||::|||::|||||::|:||:::|||||||||
Db    607 CACTGTGAAGATGGGCTATGCACCAATAGTTGTCACAAAGGCTATTCAAAAGGTTGCA 666
Qy    230 LysSerLeuLysLeuThrLeuThrCysLysHisGlnLeuValAlaArgAspSerCysLysAla 249
          |||  |||||||||||::||||||||||::||:::||||||::::||||:::||||||
Db    667 AAATATCTGAAGAAGAGATGCTATCCTGTGAACATGAACTTCTTAAAAAGGTTGCAAAGCT 726
Qy    250 SerCysAsnCysSerAsnSerIleTyr 258
          ||:::|||:::||::||||::::
Db    727 ACATGCCTCTGTGAAGGCAAAATTCAC 753
```

```
RESULT 14
MUSCRISPA

LOCUS       MUSCRISPA               1403 bp    mRNA    linear    ROD 26-JUL-1993
DEFINITION  Mouse cysteine-rich secretory protein-1 mRNA, complete cds.
ACCESSION   L05559
VERSION     L05559.1  GI:309190
KEYWORDS    cysteine-rich protein; cysteine-rich secretory protein-1; secretory
            protein.
SOURCE      Mus musculus epididymis cDNA to mRNA.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1403)
AUTHORS     Haendler,B., Kratzschmar,J., Theuring,F. and Schleuning,W.D.
TITLE       Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/AEG)
            and the novel related CRISP-3 are expressed under androgen control
            in the mouse salivary gland
JOURNAL     Endocrinology 133 (1), 192-198 (1993)
MEDLINE     93307144
PUBMED      8319566
FEATURES             Location/Qualifiers
     source          1..1403
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /tissue_type="epididymis"
     CDS             26..760
                     /function="sperm-coating protein"
                     /standard_name="CRISP-1"
                     /note="N-glycosylation site (aa): (145)"
                     /codon_start=1
                     /product="cysteine-rich secretory protein-1"
                     /db_xref="GI:309191"
                     /protein_id="AAA37460.1"
                     /translation="MALMLVLFLAAVLPPSLLQDSSQENRLEKLSTTKMSVQEEIVS
                     KHNQLRRMVSPSGSDLLKEMWNYDAQVNAQQWADKCTFSHSPIELRTTNLRCGENLFM
```

BASE COUNT   457 a   265 c   239 g   442 t
ORIGIN

SSYLASWSSAIOGWYNEYKDLTYDVGPKQPDSVVGHYTQVVWNSTFQVACGVAECPKN
PLRYYVCHYCPVGNYQGRLYTPYTAGEPCASCPDHCEDGLCTNSCGHEDKYTNCKYL
KKMLSCEHELLKKGCKATCLCEGKIH"

US-09-698-781-3 (1-258) x MUSCRISPA (1-1403)

Alignment Scores:
Pred. No.:
Score: 5.91e-69
Percent Similarity: 70.68%
Best Local Similarity: 54.62%
Query Match: 53.80%
DB: 10

                        Length: 1403
                        Matches: 136
                        Conservative: 40
                        Mismatches: 72
                        Indels: 1
                        Gaps: 1

Qy   10 GluThrThrAlaMetThrLeuPheProValLeuLeuPheLeuValAlaLeuGlyLeuLeuPro 29
          ||| ::::||||||||||||||||||||||||| |||
Db   14 GAAGCCAGCACCATGGCATTAATGCTTGTGCCTGTTCTTCTTGGCTACTGCCACCA 73

Qy   30 SerPheProAlaAsnGluAspLysAspProAlaPheThrAlaLeuLeuThrThrGlnThr 49
          ||| ::::
Db   74 TCCCTTCTTCAAGATAGCTCTCAGGAAAATCGTCTGAGAAACTTTCAAACCACTAAAATG 133

Qy   50 GlnValGlnArgGluIleValAlaAsnLysHisAsnGlyLeuArgArgAlaLeuSerProPro 69
          ||||||| |||||||||||||||||| ||||||
Db   134 TCAGTCCAAGAAGAGATTGTAAGCAAGGCACACCAATTGAGAGAGGTTTGTCCATCT 193

Qy   70 AlaAlaArgAsnMetLeuLysMetGluTrpAsnLysGluAlaAlaAlaAsnAlaGlnLysTrp 89
          |||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   194 GGCAGTGACTTACTAAAAATGGAATGGAACAAGATGCTCAAGTAGAAGCTCAGCAATGG 253

Qy   90 AlaAsnGlnCysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLysCys 109
          ||||||| ||||| || ||::::||||| |||
Db   254 GCAGACAAGGTCAGTACATTCAGTCACAGTCCTATAGAACTCAGGACAACTAATTTAAGATGT 313

Qy   110 GlyGluAsnLeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSerTrp 129
          |||||||||| ||||||||||||||||| ::||||||||||::|||
Db   314 GGGGAGAATTTGTTCATGTCATCTACCTTGCATCAACGGTCTTCTTGCAATCCAAGGATGG 373

Qy   130 PheAspGluTyrAsnAspPheAspPheGlyValGlyProLysThrProAsnAlaValVal 149
          |:|||||||| |||||||||||||||||||||||||||||||
Db   374 TATAATGAATACAAAGATCTTACATATGATGTTGGCCCAAAGCAACCTGAATAGTGTGGTT 433

Qy   150 GlyHisTyrThrGlnValValIleTrpTyrSerSerTyrLeuValGlyCysGlyAsnAlaTyr 169
          ||||||||||||||||||||||||||| |||
Db   434 GGACATTATACTCAGGTTGTTGGGAACTCAACTTCCAAGTTGCATGTGGAGTTGCTAGA 493

Qy   170 CysProAsnGlnLysValLeuLysTyrTyrThrYrValCysGlnThrYrCysProAlaGlyAsn 189
          ||||||| ||||||||||| |||||||:|||
Db   494 TGCCCT---AAAAATCCACTGAGATACTATATATTGTTCACTATTGTCCTGTTGGCAAT 550

Qy   190 TrpAlaAsnArgLeuTyrValProTyrGluGlnGlyAlaProCysAlaSerCysProAsp 209
          ::: |||||||||||||||:||
Db   551 TATCAAGGAAGGCTATACACACCTTACACTGCAGGAGAACCCGTGTCCAGTTGTCCTGAT 610

Qy   210 AsnCysGlyAspAspGlyLeuCysThrAsnGlyCysLysTyrGluAspLeuTyrSerAsnCys 229
          ||::::||||||||||||| ||| :::||||||
Db   611 CACTCTGAAGATGGGCTATGCACCAATAGTTGTGACATGGAAGATAAGTATACTAACTGT 670

Qy   230 LysSerLeuLysThrLeuThrCysLysHisGlnLeuValArgAspSerCysLysSerAla 249
          ||| ||| ||| ||||||
Db   671 AAATATCTGAAGAGAAGGATGCTATCCGTGAACATGACTTCTTAAAAAAGGTTGCAAAGCT 730

Qy   250 SerCysSerAsnCysSerAsnSerIleTyr 258
          ::::||| ||
Db   731 ACATGGCTCTGTGAAGGCAAAATTCAC 757

LOCUS   BC011150   1445 bp   mRNA   linear   ROD 07-AUG-2002
DEFINITION   Mus musculus, Similar to acidic epididymal glycoprotein 1, clone
             MGC:18489 IMAGE:4165956, mRNA, complete cds.
ACCESSION   BC011150
VERSION   BC011150.1 GI:15029853
KEYWORDS   MGC.
SOURCE   house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1445)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Baylor College of Medicine Human Genome
             Sequencing Center
             Center code: BCM-HGSC
             Web site: http://www.hgsc.bcm.tmc.edu/cdna/
             Contact: amg@bcm.tmc.edu
             Gunaratne, P.H.; Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
             Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
             Richards, S., Gibbs, R.A.
             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAK Plate: 24 Row: e Column: 21
             This clone was selected for full length sequencing because it
             passed the following selection criteria: matched mRNA gi: 6753001.
FEATURES   Location/Qualifiers
     source   1..1445
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FYB/N"
                     /clone="MGC:18489 IMAGE:4165956"
                     /tissue_type="Salivary gland, 10 week old female mouse"
                     /clone_lib="NCI_CGAP_SG2"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS   39..773
                     /codon_start=1
                     /product="Similar to acidic epididymal glycoprotein 1"
                     /protein_id="AAH11150.1"
                     /db_xref="GI:15029854"
                     /translation="MALMLVLFFLAAVLPPSLLQDSSQGENRLEKLSTTKMSVPEEIVS
                     KHNQLRRMVSPSGSDLLKMEWNRYDAQVNAQQWRADKCFSHSPIELRJTTNLRCGGENSEM
                     SSYLASWSSAIOGWYNEYKDLTYDVGPKQPDSVVGHYTQVVWNSTFQVACGVAECPKN
                     PLRYYVCHYCPVGNYQGRLYTPYTAGEPCASCPDHCEDGLCTNSCGHEDKYTNCKYL
                     KKMLSCEHELLKKGCKATCLCEGKIH"

BASE COUNT   487 a   273 c   241 g   444 t
ORIGIN

US-09-698-781-3 (1-258) x BC011150 (1-1445)

Alignment Scores:
Pred. No.:
Score: 1.02e-67
Percent Similarity: 69.88%
Best Local Similarity: 53.82%
Query Match: 52.96%
DB: 10

                        Length: 1445
                        Matches: 134
                        Conservative: 40
                        Mismatches: 74
                        Indels: 1
                        Gaps: 1

Qy   10 GluThrThrAlaMetThrLeuPheProValLeuLeuPheLeuValAlaLeuGlyLeuLeuPro 29
          ||| ::::||||||||||||||||||||||||| |||
Db   27 GAAGCCAGCACCATGGCATTAATGCTTGTGCTTTTCTTCTTGGCTACTGCCACCA 86

Qy   30 SerPheProAlaAsnGluAspLysAspProAlaPheThrAlaLeuLeuThrThrGlnThr 49
          ||| ::::

```
Db   87  TCCCTTCTTCAAGATAGCTCTCCAGGAAAATCGACTTGAGAAACTTTCAACCACTAAATG 146

Qy   50  GlnValGlnArgGluIleValAsnLysHisAsnGluLeuArgArgAlaValSerProPro 69
              |||  ||||       ::||||||||| :|||||||||||  :||| | ||||||| :
Db  147  TCAGTCCCAGAAGAGAGATTGTAAGCAAGCACAACCAATTGAGACGAGATGGTTTCCATCT 206

Qy   70  AlaArgAsnMetLeuLysMetGluTrpAsnLysGluAlaAlaAlaAsnAlaGlnLysTrp 89
             :|||:|||||||||||  ||||||||||||| ||||||| :||||:||||:|||||||
Db  207  GGCAGTGACTTACTAAAAATGGAAATGGAACTATAGATGCTCAAGTGAATGCTCAGCAATGG 266

Qy   90  AlaAsnGlnCysAsnTyrArgHisSerAsnProLysAspArgMetThrSerLeuLysCys 109
             |||:::|||  ||| ::  ||||||| |||  ||| ||||::|::|||||  ::|||
Db  267  GCAGACAAAGTGTACATTCAGTCACACTCCTATAGAACTCCAGACAACTAAATTAAGGATGT 326

Qy  110  GlyGluAsnLeuTyrMetSerSerAlaProSerSerTrpSerGlnAlaIleGlnSerTrp 129
             |||||||||||| ||||||||:  ::||||||||||| |||||| ::|||||||||
Db  327  GGGGAGAATTCTTTCATGTCAATCTTACCTTGCATCAGGTCTTCTGCAATCCAAGGATGG 386

Qy  130  PheAspGluTyrAsnAspPheGlyValGlyProLysThrProAsnAlaValVal 149
             :|||||:||||   |||      ::: |||||||||||||  ||:::||||||||||
Db  387  TATAATGAATACAAAGATCTTACATATATGATAGTTGGCCCAAAGCAACCTGATAGTGGGTC 446

Qy  150  GlyHisTyrThrGlnValValTrpTyrSerSerTyrLeuValGlyCysGlyAsnAlaTyr 169
             ||||||||||||||||||||||||| ||  |||:::::|||  |||||
Db  447  GGACATTATACTCAGGTTGTTTGGTATTCCTCACTTTCCAAGTGCATGTGGGAATGCTGAA 506

Qy  170  CysProAsnGlnLysValLeuLysTyrTyrTyrTyrValCysGlnTyrCysProAlaGlyAsn 189
             |||:||||||||||||||||||||||||||||||||| |||  |||:|||||
Db  507  TGCCCT---AAAAATCCACTGAGATACTATTATGTTTGTCACTATTGTCCTGTTGGCAAT 563

Qy  190  TrpAlaAsnArgLeuTyrValProTyrGluGlnGlyAlaProCysAlaSerCysProAsp 209
             :||        ||||||       |||||||||||||||||||:||||||||||||
Db  564  TATCAAGGAAGGCTATACACACCTTACACTGCAGGAGAACCGTGTGCCAGTTGTCCTGAT 623

Qy  210  AsnCysAspAspGlyLeuCysThrAsnGlyCysLysTyrGluAspLeuTyrSerAsnCys 229
             :::||||| ||:||||||||||||| :::||||  ||||||   ||||
Db  624  CACTGTGAAGATGGGCTATGCACCAATAGTTGTGGACATGAAGATATAAGTATACTGT 683

Qy  230  LysSerLeuLysLeuThrLeuThrCysLysHisGlnLeuValArgAspSerCysLysAla 249
             |||  ||||||||||||||||||| :::||||||||   :::::::
Db  684  AAATATCTGAAGAAGGATGCTATCCTGTGAACATGAACTTCTTAAAAAAGGTTGCAAAGCT 743

Qy  250  SerCysAsnCysSerAsnSerIleTyr 258
             :::||| |||    ||||::: 
Db  744  ACATGCCCTCTGTGAAGGCAAAATTCAC 770
```